



Open Archive Toulouse Archive Ouverte

OATAO is an open access repository that collects the work of Toulouse researchers and makes it freely available over the web where possible

This is an author's version published in: <https://oatao.univ-toulouse.fr/22143>

To cite this version:

Dousset, Bernard *Les possibilités de l'analyse des données sur les brevets : exemple en cosmétologie*. (2013) In: Veille Stratégique Scientifique & Technologique (VSST), 23 October 2013 - 25 October 2013 (Nancy, France). (Unpublished)

Any correspondence concerning this service should be sent to the repository administrator: tech-oatao@listes-diff.inp-toulouse.fr

Les possibilités de l'analyse des données sur les brevets : exemple en cosmétologie

Bernard DOUSSET (*)

bernard.dousset@irit.fr

(*) IRIT Université Paul Sabatier, Toulouse (France)

Mots clefs :

[Veille technologique](#), brevets, bibliométrie, graphes, réseaux sociaux, réseaux sémantiques, signaux faibles, géostratégie, visualisation, interactivité

Keywords:

Technology watch, patents, bibliometry, graphs, social networks, semantic networks, weak signals, geostrategy, visualization, interactivity

Palabras clave :

Vision Tecnológica, patentes, bibliometría, gráficos, redes sociales, redes semánticas, señales débiles, geoestratégico, visualización, interactividad

Résumé

Dans le cadre de nos relations suivies avec le monde de l'industrie et de la R&D, nous avons été amenés à montrer, sur un cas concret, les possibilités d'analyse des brevets et les informations stratégiques qu'il était possible d'extraire via des outils bibliométriques performants. Nous avons proposé une analyse des brevets issus de la base américaine USPTO dans le domaine de la cosmétologie et plus précisément des shampoings. Pour cela, nous avons travaillé simultanément sur les deux bases proposées par cet office à savoir la base des brevets déjà délivrés et celle des brevets publiés mais toujours en cours d'examen. Dans cette seconde base, beaucoup de brevets ne seront pas validés, mais s'y trouvent déjà les brevets qui vont être acceptés ces quatre ou cinq prochaines années. Comme nous sommes à la recherche de signaux faibles, il n'est pas possible de faire abstraction de cette seconde source, même si elle est incomplète et rédigée un peu en marge de la norme de la base initiale. Les principaux résultats présentés tournent autour des réseaux sociaux, des réseaux sémantiques (signaux forts et faibles), de la terminologie et de son évolution, de la géostratégie (côté déposants et côté inventeurs), des relations entre les déposants et les examinateurs ou les cabinets qui prennent part à la rédaction des brevets. Nous nous sommes intéressés aussi à l'antériorité afin de proposer une classification des brevets délivrés en fonction des communautés depuis lesquelles sont venues les idées innovantes.

Abstract

In the context of our relations with the world of industry and R&D, we will present, on a concrete case, the possibility of patent analysis and strategic information, which was possible to extract through the use of bibliometric tools. We have proposed an analysis of patents from the American USPTO base in the field of cosmetology more specifically on shampoos. For this, we worked simultaneously on both bases proposed by the office of patents, such as patents already granted and patent published but still under consideration. In this second base, many patents will not be validated, but some of those patents will be accepted in the next four or five years. As we are looking for weak signals, it is not possible to ignore this second source, even if it is incomplete and written somewhat outside of the norm of the initial basis. The main results presented revolve around social networks, semantic networks (strong and weak signals), terminology and evolution of the geostrategic (applicants and inventors) and the relationship between applicants and examiners or the offices involved in the drafting of patents. We are also interested in the precedence in order to provide a classification of patents based on communities from which came the innovative ideas.

Introduction

Cet article a pour but de définir une méthodologie d’analyse de la base de brevets en ligne USPTO. Nous avons choisi comme exemple les « Shampoing » afin d’illustrer cette démarche sur un cas concret, qui peut servir de cadre pour d’autres analyses similaires qui s’intéressent à un domaine technologique bien ciblé. Nous avons travaillé simultanément sur la base des brevets délivrés et sur la base des brevets en cours d’examen (déposés et publiés mais nom encore délivrés). Sur cette dernière base, il manque certaines informations comme l’antériorité, l’examineur, bien entendu la date de délivrance et bien souvent l’identité même du déposant. Le but poursuivi est multiple :

- Démontrer les capacités d’analyse des brevets
- Analyser les brevets et les demandes en cours
- Produire des statistiques descriptives
- Localiser les recherches dans le temps et l’espace
- Etudier les communautés : inventeurs/déposants
- Rechercher les signaux forts (thèmes récurrents)
- Détecter les signaux faibles (thèmes émergents)
- Détailler et investiguer à la demande

1 Extraction des brevets et reformatage

1.1 Méthode utilisée

La préparation de ce type d’étude impose de respecter un certain nombre d’étapes qui peuvent faire appel à des outils maison ou en open source :

- Mettre au point une équation de recherche
- Extraire les URL de tous les brevets
- Télécharger chaque brevet au format html (utilisation d’un robot)
- Reformater le corpus et compléter le balisage (programme perl maison)
- Extraire le dictionnaire de chaque champ
- Corriger les informations (inventeurs, pays, états, codes)
- Générer les multi-termes issus du texte libre
- Calculer les matrices de cooccurrences 2D et 3D nécessaires à l’étude
- Fusionner les brevets et les demandes en cours

1.2 Les métadonnées pour la base USPTO

Afin d’adapter le fonctionnement de nous outils d’extraction au format de la base USPTO, nous devons décrire ce format par des métadonnées qui prennent en compte les champs balisés, leur structure, les séparateurs utilisés et éventuellement les hiérarchies ou les contenus complexes (adresses, sources, texte libre). Ces métadonnées sont décrites grâce à un tableau comportant 5 colonnes : le nom explicite du champ logique, son abréviation, le nom du champ physique auquel il est rattaché, un booléen pour dire s’il sera utilisé ou pas, une liste de séparateur permettant le hachage du contenu.

USP:					Date_Depot	DD	DP:	True	," ORD2"
Date_Pub	DP	Usp:	True	,b"*ORD0"	Famille	FA	FI:	True	"
Numero	NO	USP:	True	("ORD1"	PCT_Filed	Dd	Dd:	False	"
Abstract	AB	AB:	True	,",",":"/""(")"b""	PCT_No	Np	Np:	False	"
Titre	TI	TI:	True	,",",":"/""(")"b""	Date_Issue	Di	Di:	False	"
Inventeur	AU	AU:	True	(",)b ").)"	Num_Ref	Nr	Nr:	False	"
Pays_Inventeur	Pa	AU:	True)("	Date_Ref	Dr	Dr:	False	"
Etat_Inventeur	Us	AU:	True)("	Related_US_Pat_Do	RU	RU:	False	"
Deposant	CO	CO:	True	(",)b ").)"	Ref_Cited	RC	RC:	False	"
Pays_Deposant	PA	CO:	True	(")"	For_App_Prio_Data	FA	FA:	False	"
Etat_Deposant	US	CO:	True	(")"	Class_US	UC	UC:	True	;" "
Num_Appl	NA	AN:	True	"	Class_WO	IC	IC:	True	;" "&(":"

Gov_Interest	GI	GI:	False	"
Champ_Rech	FS	FS:	True	b"
US_Patent_Doc	UD	US:	True	"
For_Pat_Doc	FO	FO:	False	"
Other_Ref	AR	AR:	False	"

Examineur	EX	PE:	True	"
Ass_Exam	AE	AE:	True	"
Agent	AG	AA:	True	"ORD1"
Description	DE	DE:	False	"
Claims	CL	CL:	False	"

1.3 Corrections morphologiques

La différence entre une base de données et une base documentaire est principalement due à la variabilité des formes terminologiques utilisées pour désigner un même item (inventeur, déposant, examinateur, agent, pays, état, ville, mot-clé, code de classification, sans compter les termes libres). Nous devons donc, avant toute démarche statistique, homogénéiser le vocabulaire, les entités nommées, les codes et même les formats des dates. Ci-dessous des extraits des dictionnaires de correspondances morphologiques issus du corpus étudié. Ils concernent les noms des inventeurs et ceux des déposants, un travail analogue devant être réalisé pour les pays et les états américains très maladroitement codés sur 2 lettres dans cette base (CA est équivalent à Californie, lorsqu'il s'agit des US, sinon à Canada).

- **Cas des inventeurs :**

BIRTWISTLE; DAVID HOWARD	BIRTWISTLE; DAVID H.
BROWN; WILLIAM J.	BROWN; WILLIAM
CHUN; HO-MING	CHUN; HO M.
DOWELL; TERESA JOLANTA	DOWELL; TERESA J.
HALLORAN; DANIEL JOSEPH	HALLORAN; DANIEL J.
HITCHEN; DAVID ANDREW	HITCHEN; DAVID A.
INMAN; EVERETT J.	INMAN; EVERETT
INMAN; EVERETT JUNIOR	INMAN; EVERETT
KORKIS; GEORGE NOEL	KORKIS; GEORGE N.

LOCKWOOD; M. OLENE
MURRAY; ANDREW MALCOLM
NEWELL; GERALD PATRICK
PATEL; AMRIT M.
PYLES; DANIEL RAYMOND
ROBBINS; CLARENCE R.
TAN-WALKER; RUBY LOO
WELLS; ROBERT LEE

LOCKWOOD; MARY OLENE
MURRAY; ANDREW M.
NEWELL; GERALD P.
PATEL; AMRIT
PYLES; DANIEL R.
ROBBINS; CLARENCE
BICK TAN-WALKER; RUBY LOO
WELLS; ROBERT L.

- **Cas des déposants :**

AKZO NOBEL NV
ALBERTO-CULVER COMPANY
CHESEBROUGH-POND'S USA CO., DIVISION OF CONOPCO, INC.
PROCTER & GAMBLE COMPANY
PROCTER & GAMBLE
PROCTOR & GAMBLE COMPANY
SOCIETE ANONYME DITE: L'OREAL
TAKARA BELMONT CORPORATION
THE PROCTER & GAMBLE COMPANY
THE PROCTER & GAMBLE
THE PROCTOR & GAMBLE COMPANY
THE PROCTOR & GAMBLE COMPANY

AKZO NOBEL NV.
ALBERTO CULVER COMPANY
CHESEBROUGH-POND'S USA CO.
THE PROCTER & GAMBLE CO.
THE PROCTER & GAMBLE CO.
THE PROCTER & GAMBLE CO.
L'OREAL
TAKARA BELMONT CORP.
THE PROCTER & GAMBLE CO.
THE PROCTER & GAMBLE CO.
THE PROCTER & GAMBLE CO.
THE PROCTER & GAMBLE CO.

1.4 Décodage des classifications internationales et américaines.

Dans le cas des codes de classification, seul le code est présent dans la base, son libellé doit être récupéré dans une autre partie du site d'USPTO ou sur le Web. Or il est plus qu'utile que cette information sémantique soit disponible en clair dans l'analyse. Ici encore nous utilisons le même mécanisme qui remplace chaque code par son libellé. Nous sommes obligés de faire cette transformation pour les deux types de codes présents : les classes internationales (IC) et les classes américaines (UC).

- **Classes Internationales (IC)**

A61Q 5/00 Préparations pour les soins des cheveux
A61Q 5/02 Préparations pour le lavage des cheveux
A61Q 5/04 Préparations pour onduler de façon permanente ou décrepser les cheveux
A61Q 5/06 Préparations pour mettre les cheveux en forme, p.ex. pour mettre en forme ou colorer temporairement

- **Classes US (UC)**

510 CLEANING COMPOSITIONS FOR SOLID SURFACES, AUXILIARY COMPOSITIONS THEREFOR, OR PROCESSES OF PREPARING THE COMPOSITIONS
510/119 For human scalp hair, scalp, or wig (e.g., shampoo, etc.)
510/121 Steroid or carbohydrate gum component (e.g., lanolin, guar gum, etc.)
510/122 Silicon, boron, or phosphorus containing organic component
510/123 Amine oxide, quaternary, or zwitterion nitrogen component (e.g., betaine, sultaine, etc.)

2 Statistiques descriptives

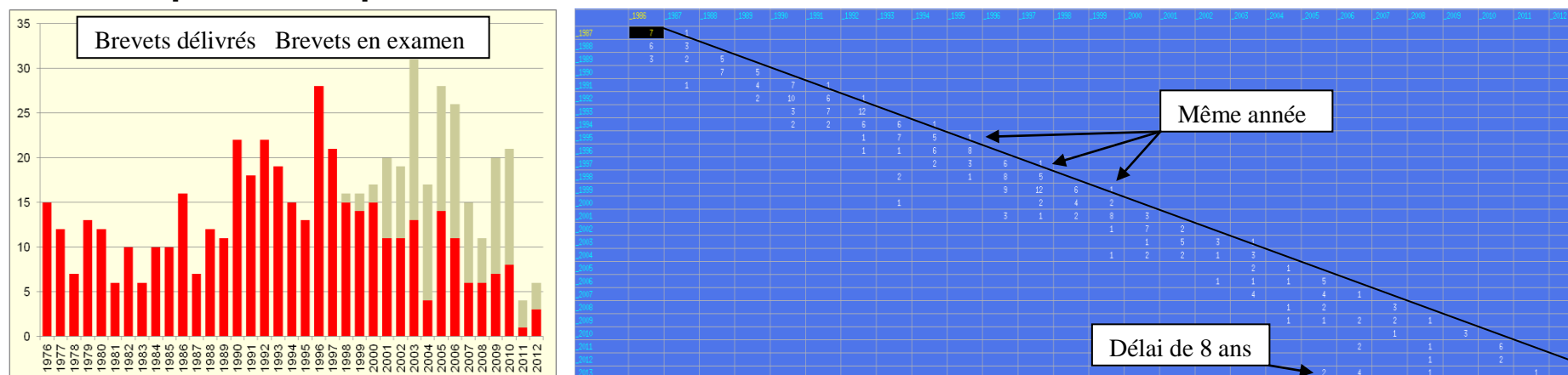


Figure 1 : Evolution du nombre de brevets sur les shampoings et évolution de l'écart entre les date de dépôt et de délivrance

Ci-dessus, les dates de dépôt des brevets déjà délivrés et de ceux qui sont publiés sur la base USPTO, mais toujours en examen. Certains ne sortiront jamais, mais ils représentent plus de 50% des brevets déposés depuis 15 ans et plus de 75% de ceux déposés depuis 10 ans. A droite, le temps de latence des brevets délivrés. Nous pouvons remarquer que certains d'entre eux sont délivrés dans l'année !

Ci-dessous, la répartition des plus fortes fréquences des classifications internationales et américaines des brevets, les distributions sont différentes un premier indice sur leur non équivalence.

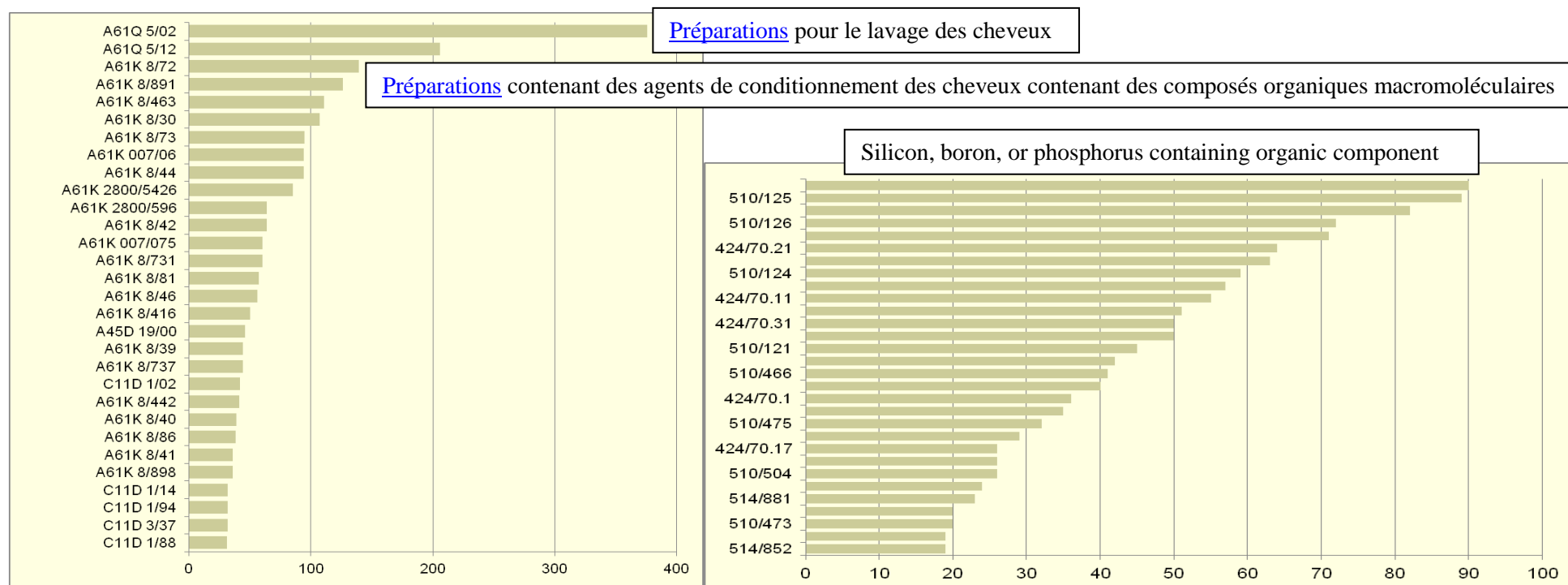


Figure 2: Fréquences des classifications internationales et américaines.

3 Détection des réseaux sociaux

3.1 Cas statique

Le principal réseau social que nous pouvons étudier est celui des co-inventeurs de brevets. Il nous renseigne sur les plus grosses équipes de R&D ainsi que sur leurs liens éventuels. Comme il s'agit le plus souvent de recherche privée, les connexions ne sont pas très nombreuses, mais elles peuvent nous permettre de détecter des ilots particulièrement actifs. Dans la figure ci-dessous, nous pouvons noter la présence de plusieurs groupes dont un très actif, la taille des sommets et proportionnelle au nombre de brevets déposés, la valeur des arêtes indiquée en rouge représente le nombre de co-dépôts entre deux inventeur. Les inventeurs isolés ont été masqués, les équipes de 2 à 4 personnes sont les plus nombreuses, mais elles n'interviennent le plus souvent que pour un ou deux brevets. Certains brevets font apparaître des équipes plus importantes jusqu'à une dizaine d'inventeurs.

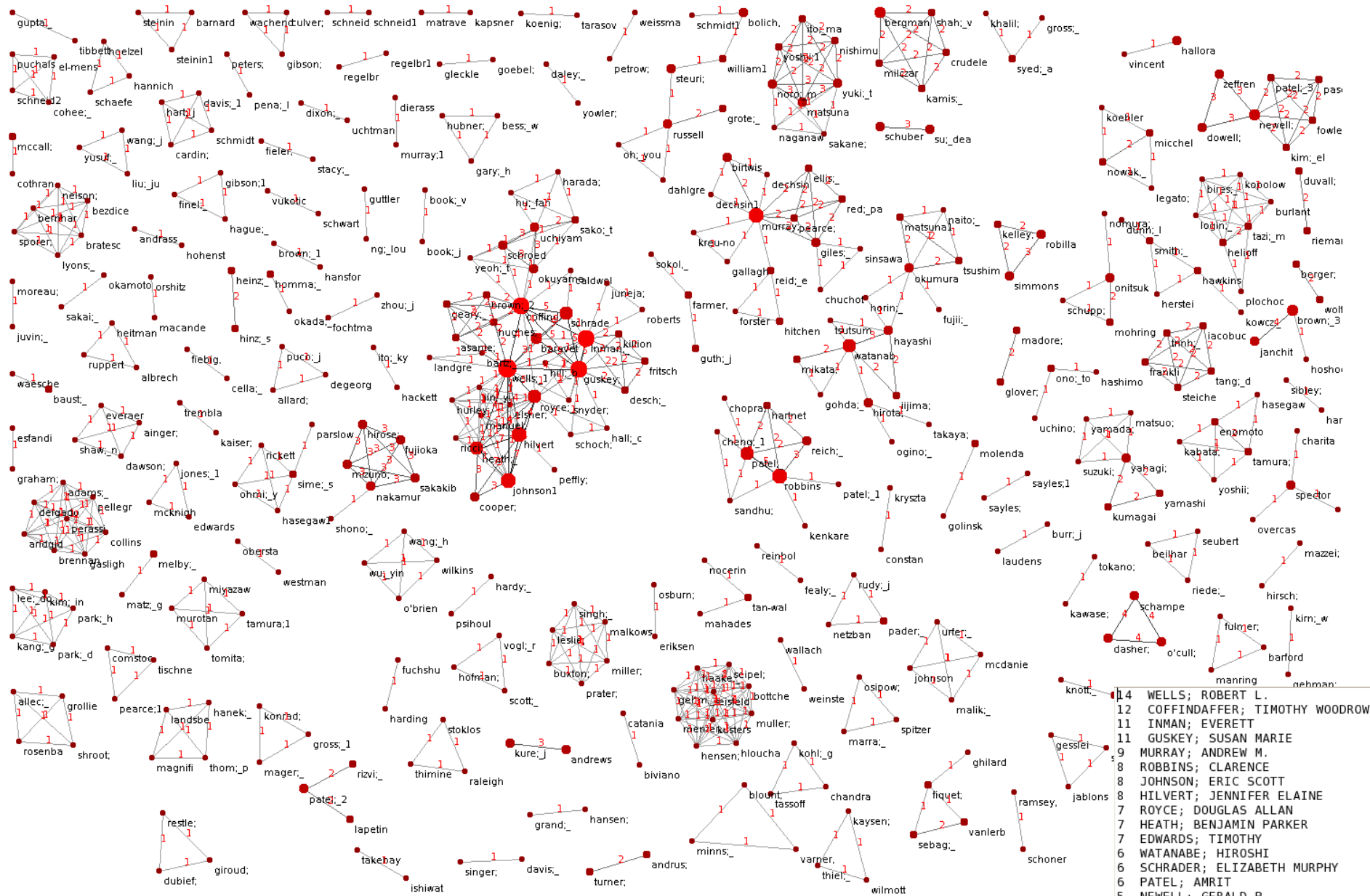


Figure 3 : Réseau social des inventeurs et fréquences des plus actifs.

Les trois principaux clusters ont été extraits du graphe précédent, afin de mieux en appréhender l'organisation et de détecter les acteurs structurants. Ci-dessous, positions stratégiques d'Inman, Murray et Hayashi.

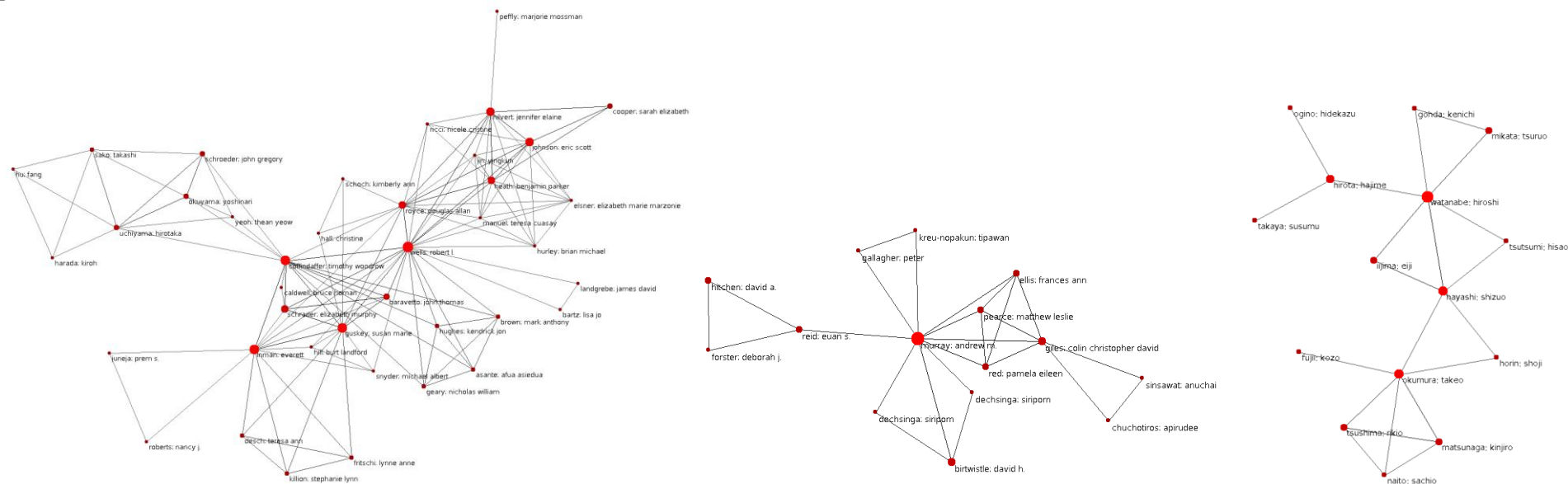


Figure 4 : Organisation des trois principales équipes.

3.2 Cas dynamique

Une analyse en dynamique va nous permettre de situer dans le temps l'activité de chaque équipe. Pour cela, nous décomposons la matrice du graphe en plusieurs plans associés chacun à une période de dépôt. Nous avons ici choisi 4 périodes que nous codons au niveau de l'histogramme représentant l'évolution de l'activité d'un sommet par une couleur allant du rouge au vert en passant par l'orange et le jaune. Afin de mieux comprendre l'évolution du réseau social, nous attirons les sommets vers des repères représentant chaque période et disposés dans le sens des aiguilles d'une montre (Repère i). Ainsi, le graphe se décante, les sommets n'appartenant qu'à une période se situent à côté d'elle, ceux qui appartiennent à deux, consécutives ou non, se trouvent entre les deux ... Au centre du graphe, nous trouvons les équipes permanentes (il y en a peu) et nous visualisons très clairement le turn-over de chacune. Les équipes qu'en rouge, orange ou jaune ne sont plus actives, celles qui ont du vert continuent à breveter, celle qui n'ont que du vert sont nouvelles sur le secteur. Nous retrouvons bien évidemment le principal cluster (Wells, Inman), nous remarquons qu'il est surtout présent dans les 2 dernières périodes, très peu dans la seconde et pas du tout dans la première. Il s'agit donc d'une équipe assez nouvelle, Inman un des pionniers est en déclin, par contre Wells a pris le relais avec trois autres inventeurs (Heath, Johnson, Hilvert) qui sont très actifs dans la dernière période. De nouveaux inventeurs apparaissent comme Cooper et Peffly, mais la moitié des inventeurs de l'équipe n'est plus active. Dans la seconde équipe, nous constatons un turn-over important avec un permanent Murray. Quant à la troisième, elle n'est plus en activité, juste présente dans la première période. La dynamique de chaque réseau peut donc être analysée finement, mais aussi l'activité globale du secteur dans lequel nous voyons apparaître un nombre impressionnant de nouvelles équipes qui pour certaines viennent de déposer plusieurs brevets. Cette approche est à compléter par l'analyse des brevets en cours

d'examen afin d'anticiper les innovations qui vont éventuellement être validées par la délivrance des dits brevets. A surveiller les nouveaux entrants et leur technologie qui peut devenir dominante à plus ou moins long terme [5], [6].

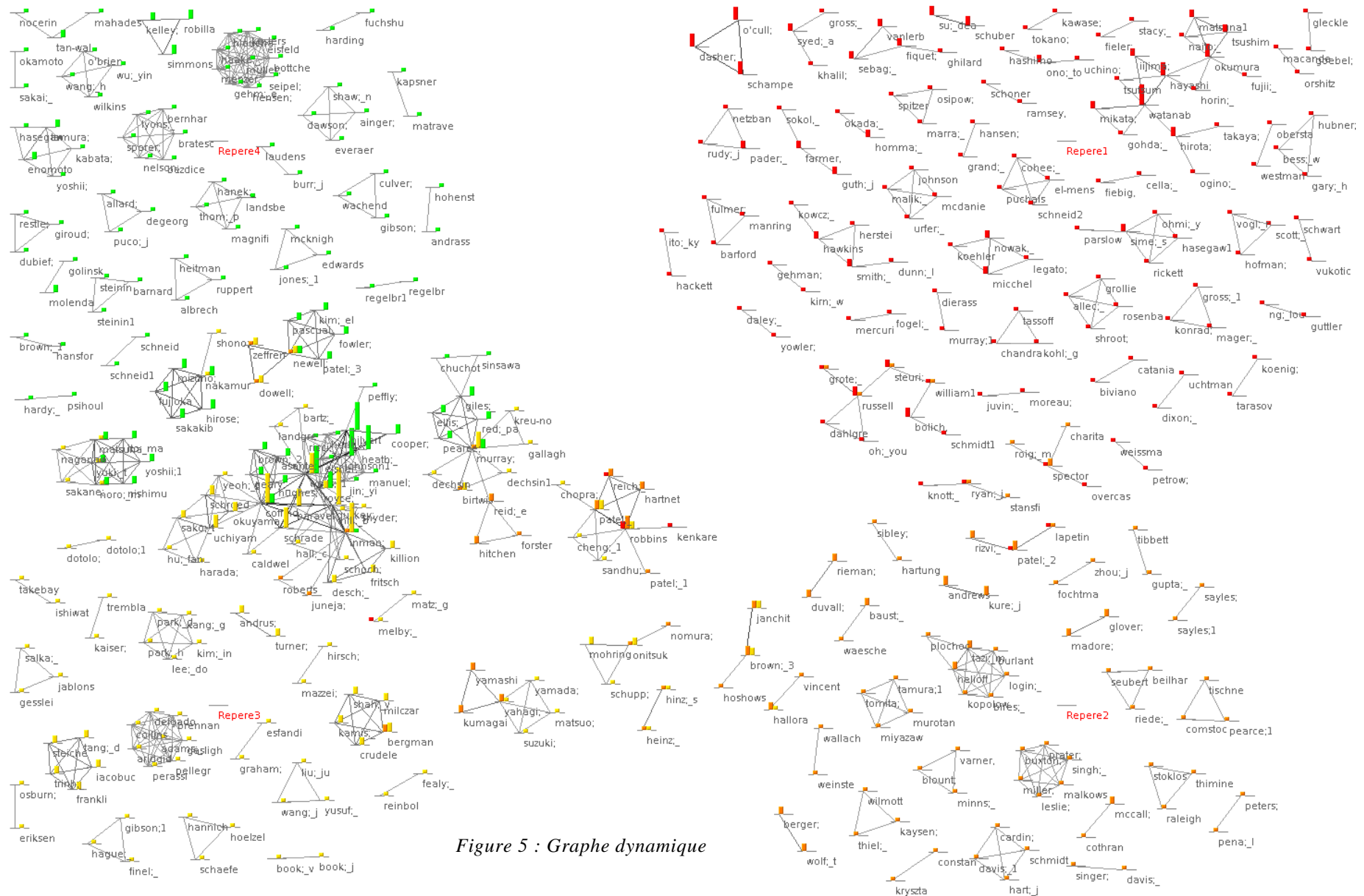


Figure 5 : Graphe dynamique

3.3 Cas biparti

3.3.1 Collaborations ou concurrences

Ces équipes peuvent être étudiées séparément en statique comme en dynamique. Elles seront croisées avec les déposants afin de faire apparaître des collaborations cachées ou des concurrences (alternances de dépôts entre plusieurs déposants via les mêmes inventeurs ou passage, sans retour, d'un ou plusieurs inventeurs d'une entreprise à l'autre) [4].

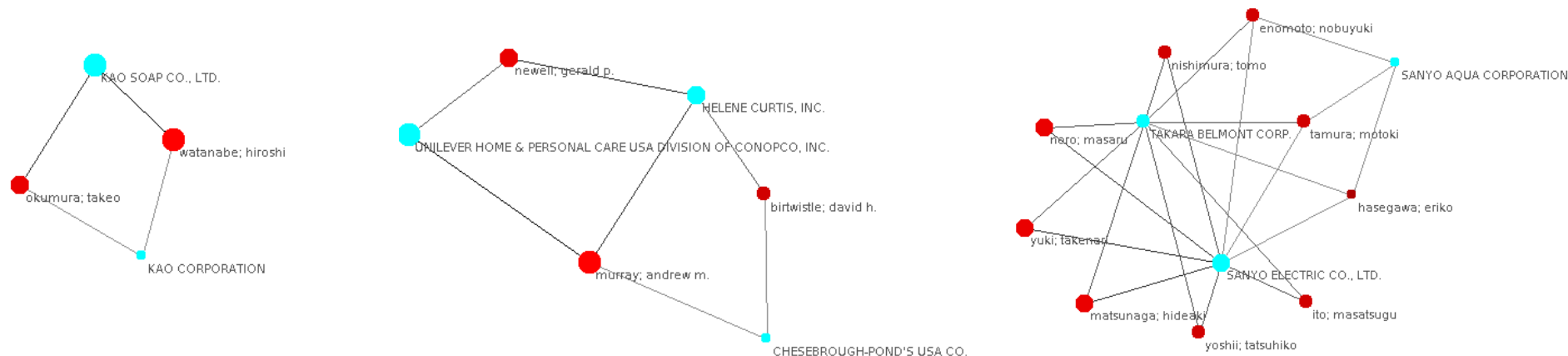


Figure 6 : graphe biparti Inventeurs(en rouge) X Déposants(en bleu)

Depuis le graphe ci-dessus, il est possible d'extraire des sous-graphes soit en utilisant la transitivité (graphe égocentrique), soit en utilisant les k-cores (graphe dont tous les sommets ont au moins k-voisins) ou par filtrage sur une mesure calculée sur les sommets comme la centralité. De tels sous-graphes permettent de repérer plus facilement certains sommets remarquables qui jouent un rôle particulier dans la topologie du graphe ou qui ont un environnement très typé. Ces fonctions de découverte de connaissances dans le relationnel offrent à l'utilisateur une aide à l'investigation qui peut s'avérer très utile quand le graphe initial est trop complexe et donc peu lisible (grand nombre de sommets, degré élevé, mal convergé, de répartition assez uniforme).

3.3.2 Liens avec les examinateurs

Les brevets sont analysés par des examinateurs d'USPTO avant d'être délivrés s'ils sont jugés recevables. Cette étape consiste notamment à juger, via l'étude d'antériorité, si l'invention est suffisamment originale, si la description est conforme et les revendications acceptables. L'examineur choisi ainsi que l'assistant examinateurs doivent être experts du domaine, ils peuvent également être désignés parce qu'ils ont l'habitude de travailler avec le déposant du brevet. Nous pouvons faire apparaître et analyser les réseaux existants entre Déposants et Examineurs par un graphe biparti qui nous renseigne sur les habitudes de travail et les liens qui s'établissent entre eux.

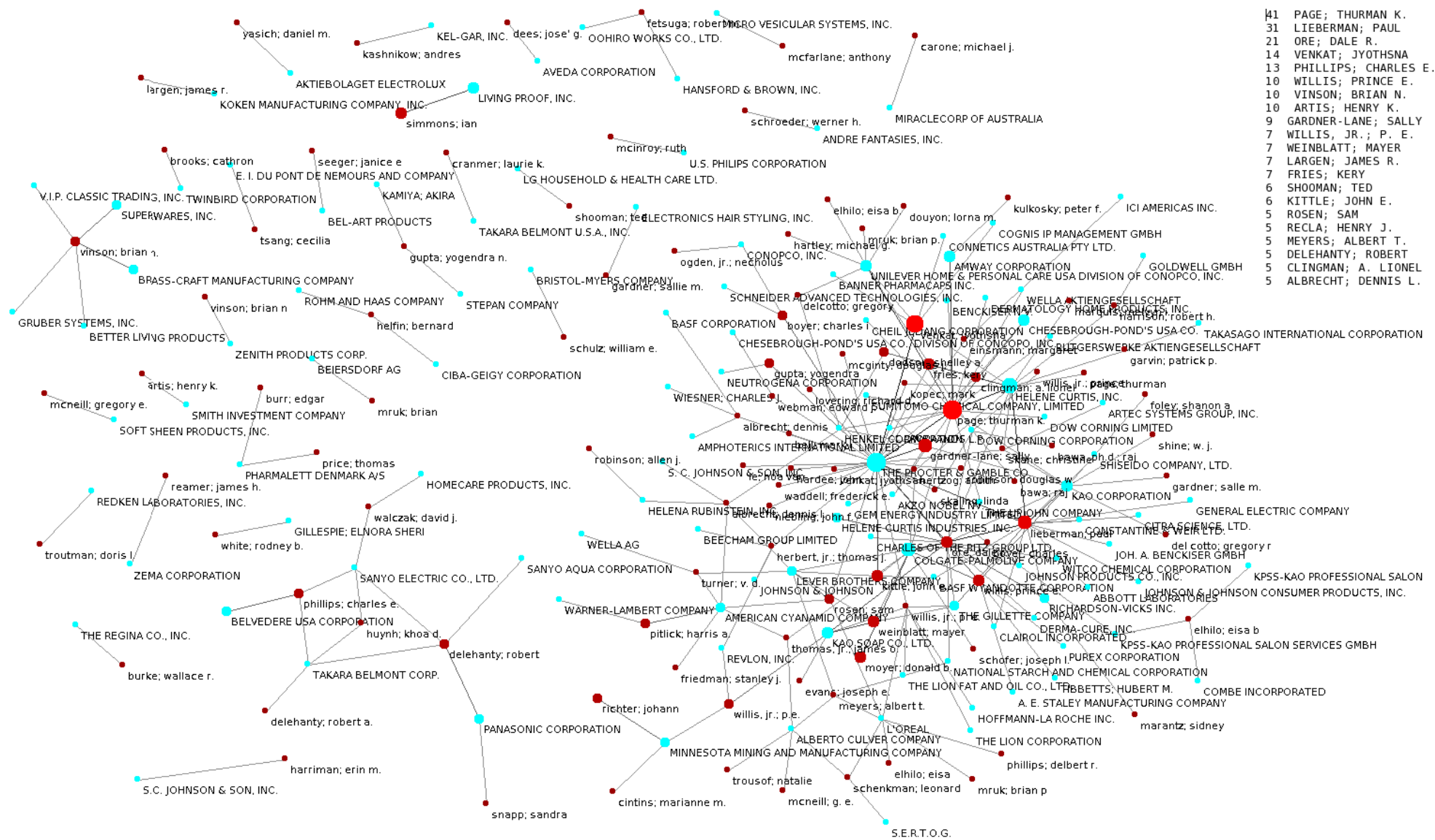


Figure 7 : graphe biparti Examineurs (en rouge) X Déposants (en bleu)

3.3.3 Liens avec les agents

Les agents interviennent au moment de la rédaction des brevets et suivent la procédure d'examen jusqu'à leur délivrance. Ils peuvent aussi s'occuper des problèmes d'extension du brevet notamment à l'international. Il est possible d'étudier le réseau des clients de ces agents, mais aussi la communauté d'agents qui travaillent pour un seul déposant ou un groupe de déposants. Certains agents changent parfois de cabinet, mais les déposants peuvent leur rester fidèles. Une étude en dynamique est alors nécessaire pour suivre ce type d'évolution.

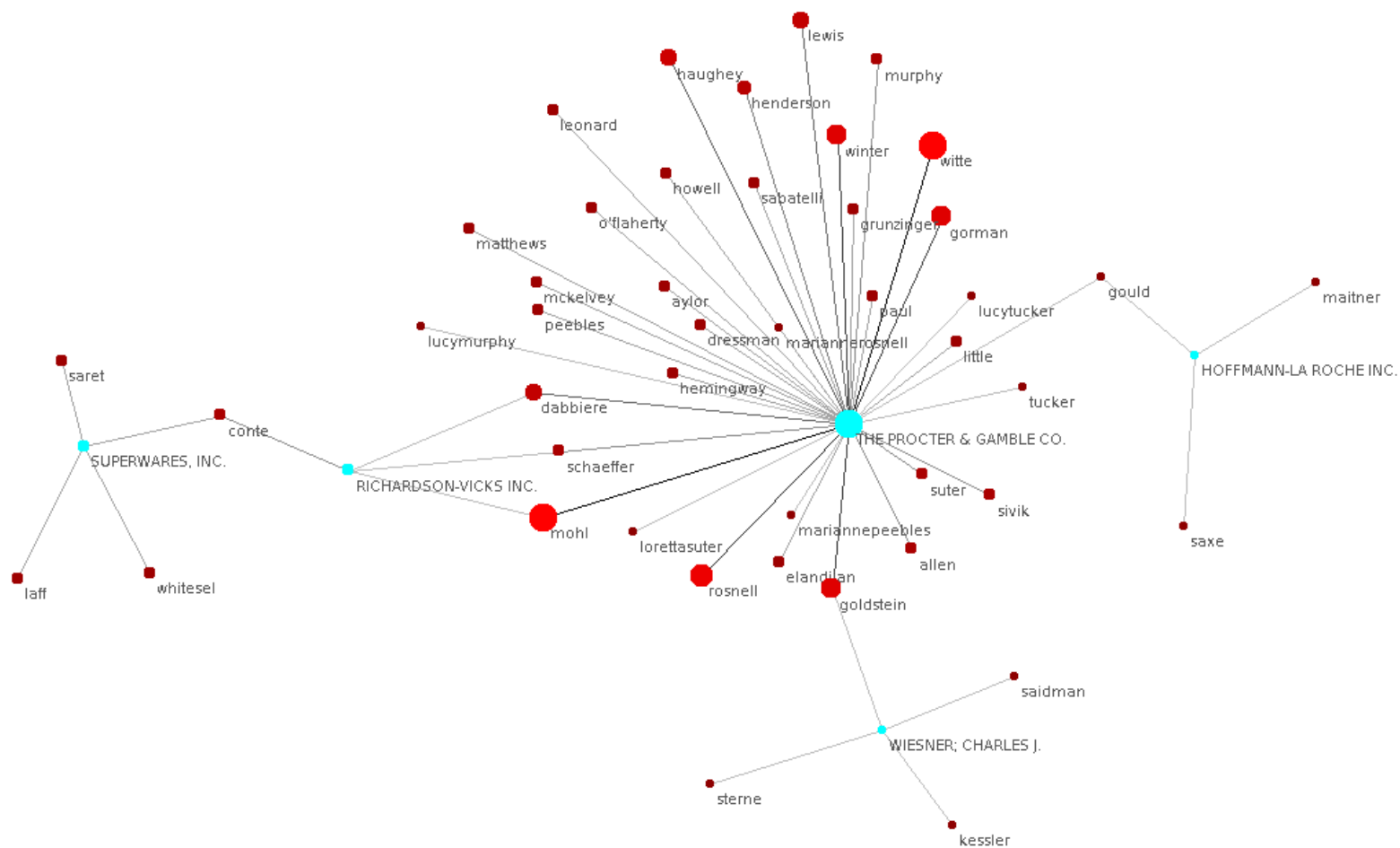


Figure 8 : graphe biparti Agents (en rouge) X Déposants (en bleu)

4 Détection des réseaux sémantiques

4.1 Codes de la classification internationale



376	A61Q	5/02	44	A61K	8/737
206	A61Q	5/12	44	A61K	8/39
139	A61K	8/72	42	C11D	1/02
126	A61K	8/891	41	A61K	8/442
111	A61K	8/463	39	A61K	8/40
107	A61K	8/30	38	A61K	8/86
95	A61K	8/73	36	A61K	8/898
94	A61K	8/44	36	A61K	8/41
94	A61K	007/06	32	C11D	3/37
85	A61K	2800/5426	32	C11D	1/94
64	A61K	8/42	32	C11D	1/14
64	A61K	2800/596	31	C11D	1/88
60	A61K	8/731	27	C11D	1/29
60	A61K	007/075	27	A61Q	5/006
57	A61K	8/81	26	C11D	1/38
56	A61K	8/46	26	A61Q	5/06
50	A61K	8/416	26	A61K	8/895
46	A45D	19/00	26	A61K	8/466

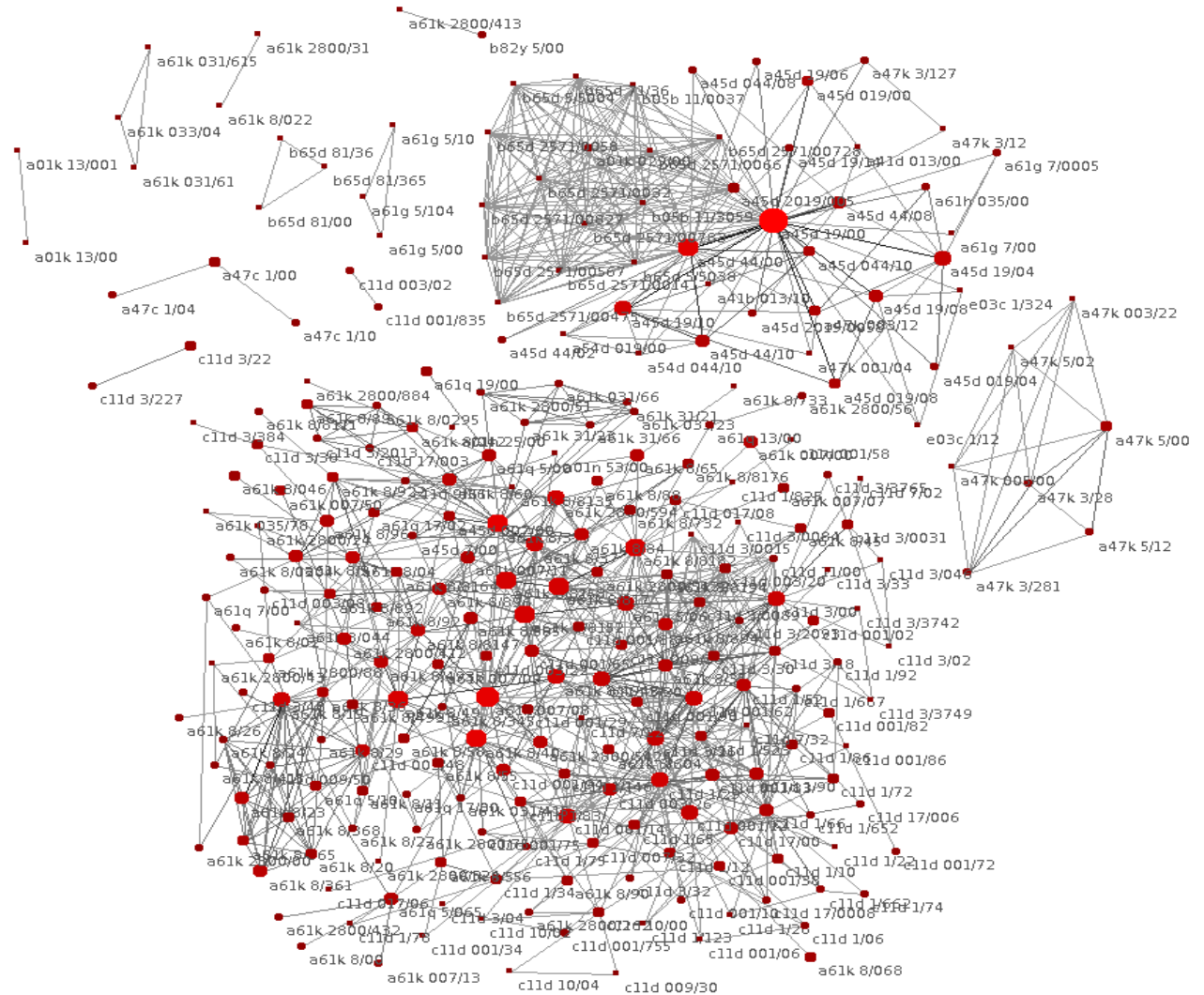


Figure 7 : clusters sémantiques (classification internationale)

4.2 Codes de la classification américaine

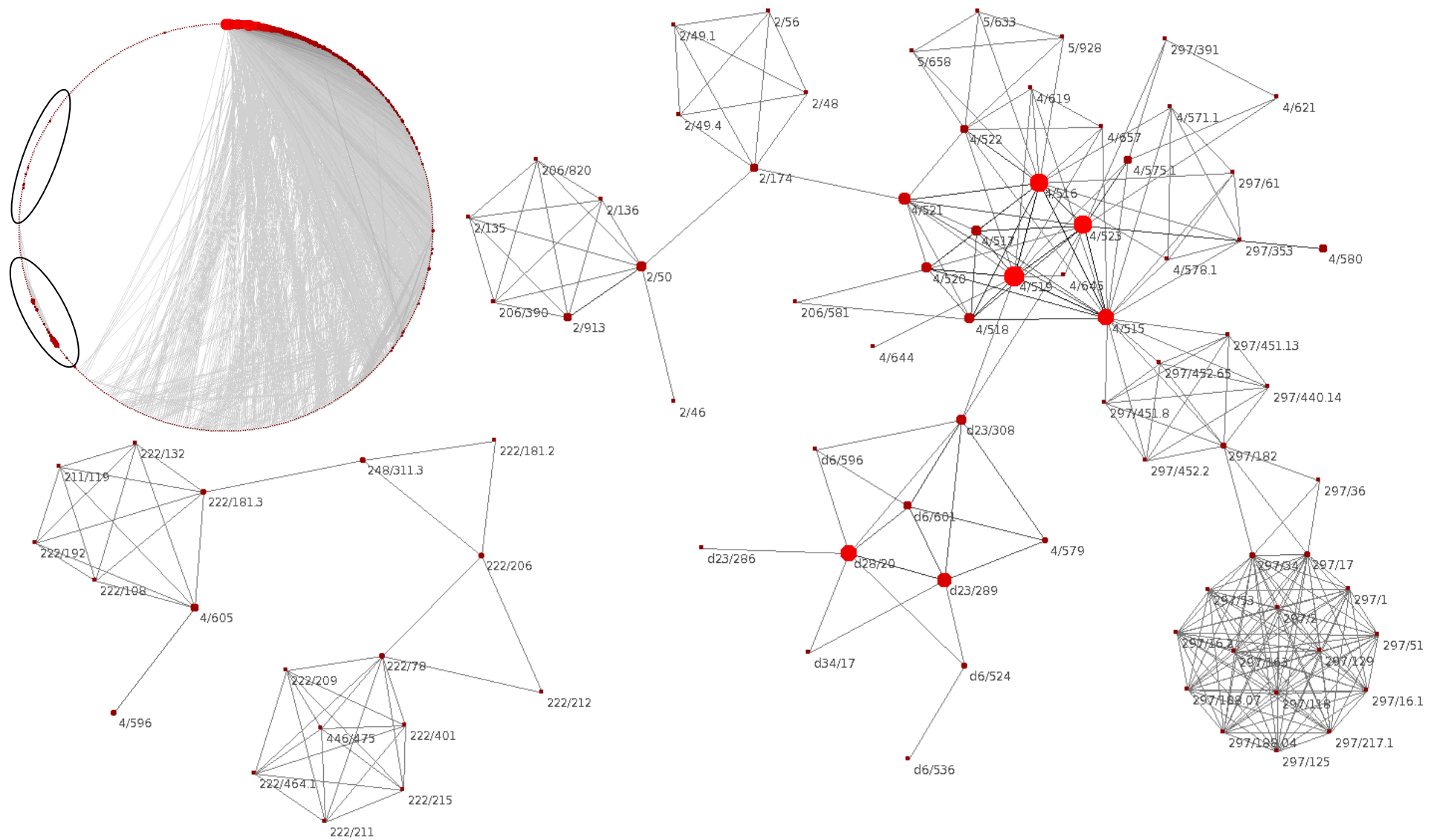


Figure 8 : Cluster sémantiques après filtrage (classification américaine)

4.3 Evolution de la terminologie

Les sujets abordés dans les brevets évoluent au cours du temps. Il est possible par une analyse factorielle des correspondances (AFC) de les situer par rapport à plusieurs périodes consécutives et de taille assez comparables. Ici, nous n'avons choisi que 4 périodes et la vue ci-contre (carte factorielle) nous permet d'extraire instantanément les termes nouveaux, ceux qui sont abandonnés ou ceux qui se maintiennent. Si le nombre de périodes est plus important, comme chacune d'entre elles est plus corrélée avec les périodes qui lui sont adjacentes, le temps s'écoule de façon continue sur la carte (effet Guttman) et il est donc assez facile de déterminer le profil d'évolution de chaque terme. Cette méthode d'analyse de l'évolution est applicable à toutes les données issues des l'analyse textuelle (mots clés, codes de classification, index, n-grams, ...) mais aussi aux acteurs du domaine (inventeurs, déposants, examinateurs, agents, ...)

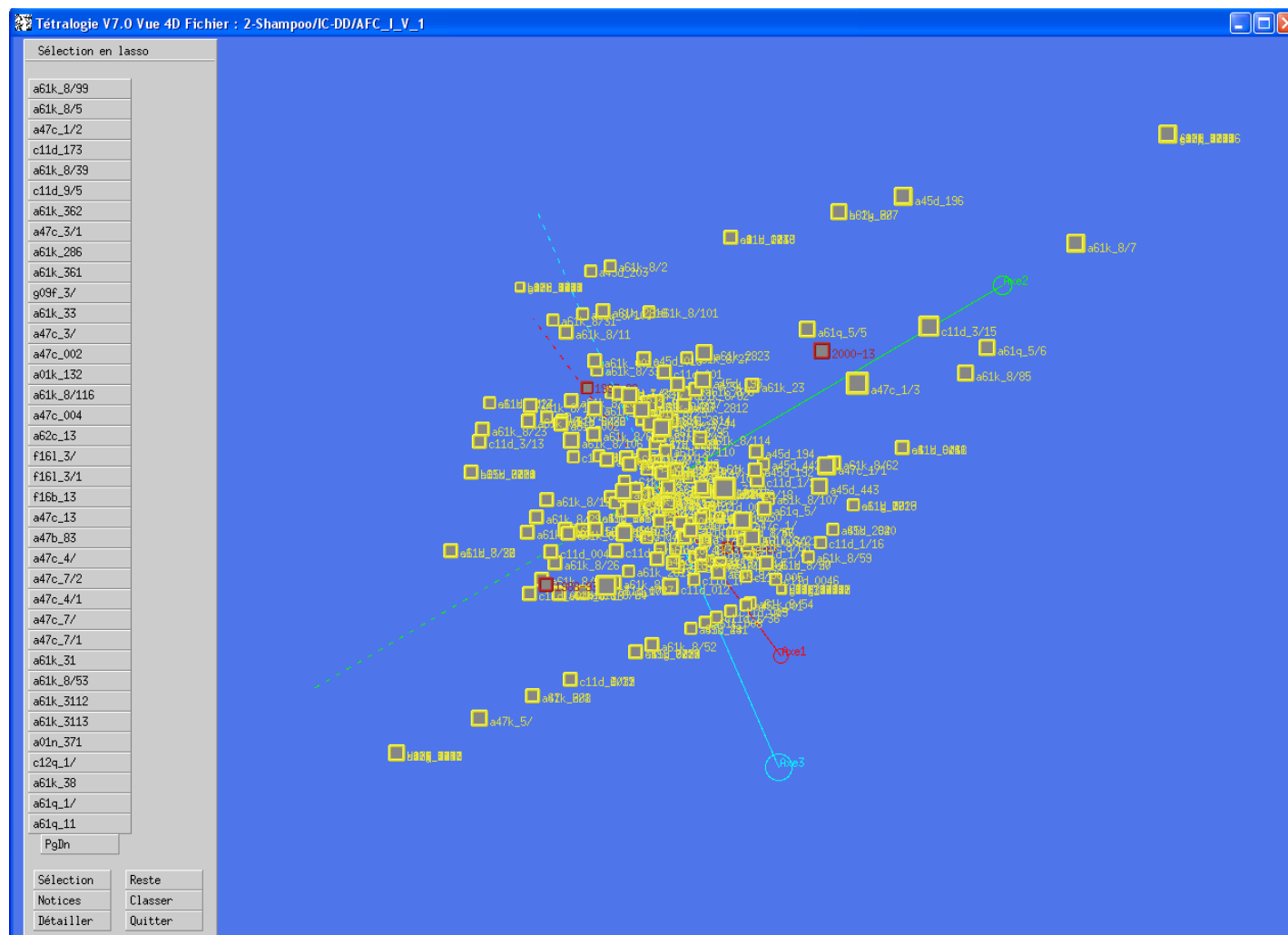
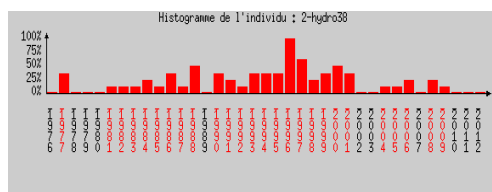


Figure 10 : Evolution du domaine (classification internationales)

4.4 Détection des multi-termes

Pour aller plus loin que les classifications (internationales et américaines), il est possible d'analyser sémantiquement le texte libre (titre, résumé, texte intégral : description et revendications). Une première approche consiste à rechercher les mots clés du domaine dans ces textes, il faut pour cela disposer de dictionnaires ou d'une ontologie de domaine. Si ce n'est pas le cas, il est toujours possible de détecter les n-grams (ou multi-termes) dans ces textes en utilisant des techniques d'analyse sémantique. Le résultat doit être filtré afin d'enlever les termes non significatifs, mais le résultat a le mérite d'être à jour car il intègre les dernières évolutions terminologiques et des néologismes souvent très pertinents et qui ne se trouvent ni dans les dictionnaires ni encore dans l'ontologie de domaine.

Voici un extrait des multi-termes issus automatiquement du texte libre des brevets.

6-AMINO BENZOMORPHOLINE
6-BIS-HYDROXYETHOXY-3
6-BRIPHOS
6-CARBON
6-CARBOXYMETHOXY-1
6-CHLORO-9-ETHY-9H-CARBAZOL-3-YL
6-COCONUTACYLAMIDO-3-TRIMETHYLAMMONIOHEXANOATE
6-DI-0-METHYL
6-DIAMINOPYRIDINE
6-DIAMINOPYRIMIDINE
6-DIAMINOTOLUENE
6-DICARBOXYLIC
6-DICHLORO-2-METHYL-1
6-DICHLORO-M-AMINOPHENOL
6-DIETHYL-PARA-PHENYLENEDIAMINE
6-DIHYDROPYRAZINE
6-DIHYDROXY-3
6-DIHYDROXY-4-METHYLPYRIDINE
6-DIHYDROXYINDOLE
6-DIHYDROXYPYRIDINE
6-DIMETHOXYPYRIDINE
6-DIMETHYL-1
6-DIMETHYL-2
6-DIMETHYL-3
6-DIMETHYL-7-OCTEN-2-OL
6-DIMETHYL-8-METHYLNONANE
6-DIMETHYLAMINO-3-METHYL-1H-INDOL-2-YL
6-DIMETHYLCYCLOHEXYL
6-DIMETHYLDODECA-2

6-DIMETHYL-PARA-PHENYLENEDIAMINE
6-DIMETHYL-P-PHENYLENEDIAMINE
6-DIMETHYL-PYRAZOLO[1
6-DIMETHYLPYRAZOLO[1
6-DIMETHYLPYRAZOLO[3
6-DIMETHYLUDECA-2
6-DINITRO-3
6-DINITRO-3-METHOXY-1-METHYL-4-TERTIARY-BUTYLBENZENE
6-DINITROINDAN
6-DIPHENYL-1
6-DI-TERT-BUTYL-4-METHYLPHENOL
6-DI-TERT-BUTYL-P-CRESOL
6-ETHOXY-1
6-ETHOXY-1-BIS
6-ETHYLDECAMETHYLENE
6-HEPTADIEN-3-ONE
6-HEXAHYDRATE
6-HEXAMETHYL-1
6-HEXAMETHYLINDAN
6-HEXANETRIOL
6-HYDROXY-2-PHENYL-3
6-HYDROXYBENZOMORPHOLINE
6-HYDROXYETHOXY-1
6-HYDROXYETHYL-1
6-HYDROXYINDOLE
6-HYDROXYINDOLINE
6-METHOXY-1
6-METHOXY-1-AMINO-3-[
6-METHOXY-5-ETHYL-1

4.5 Qui fait quoi

En croissant un champ sémantique (ici les classes internationales) avec les déposants, il devient possible d'établir une cartographie mixte qui nous renseigne sur les thèmes communs ou spécifiques, ainsi que sur la variété présente dans le domaine. Une classification sur les déposants en fonction des thèmes qu'ils abordent, nous permet de situer immédiatement les concurrences ou les espaces laissés libres.

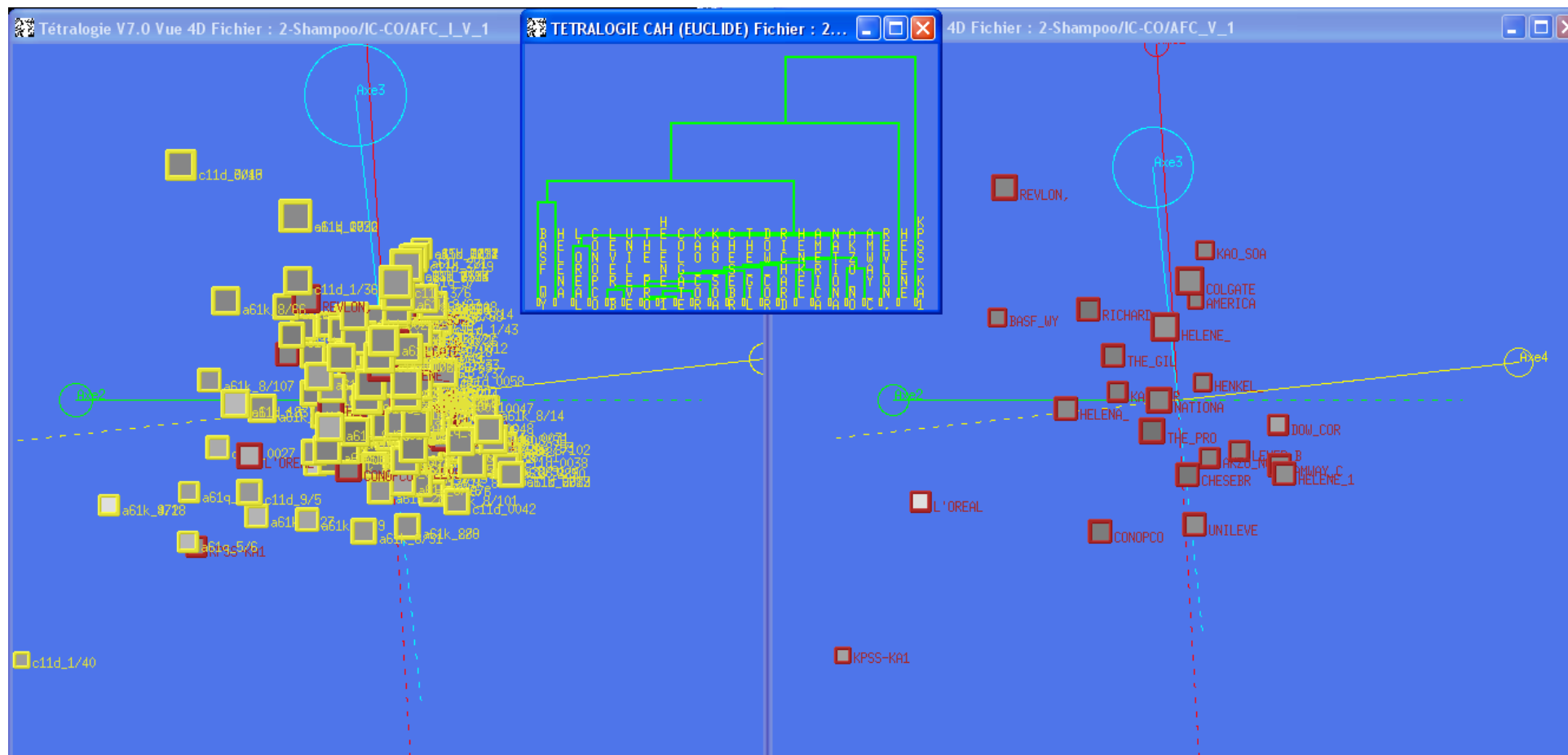


Figure 11 : Cartographie conjointe des thèmes et des déposants

5 Eléments de géostratégie

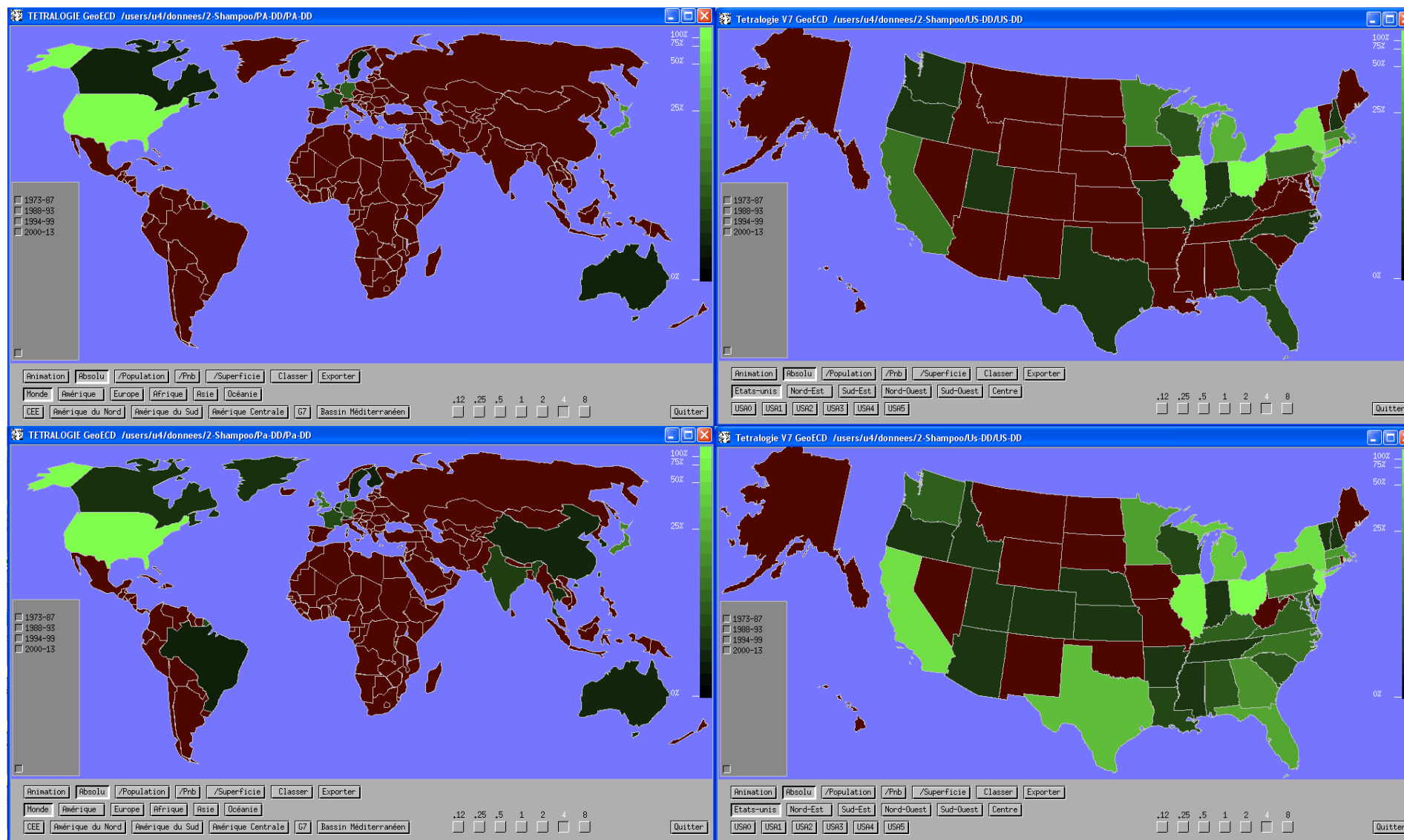


Figure 11 : Localisation des déposants (en haut) et des inventeurs (en bas) dans le monde et aux USA.

6 Conclusion

En prenant comme exemple d'analyse « *les brevets en cosmétologie dans le domaine des shampoings* », cette communication avait pour but :

- de montrer les possibilités d'analyse des bases brevets
- de prendre en compte la spécificité de chaque base
- de montrer comment compléter et corriger les informations brutes
- de détecter les réseaux sociaux
- de détecter les réseaux sémantiques
- de localiser les recherches dans le temps et l'espace
- de détecter les signaux faibles
- d'extraire des éléments stratégiques et de les mettre sous une forme assimilable.

Pour une analyse finalisée, il est nécessaire de produire en plus l'ensemble des dictionnaires extraits, des statistiques obtenues (fréquences, corrélations, cooccurrences, ...) et de donner accès à la manipulation interactive, accessible à distance [2], [10], notamment de l'ensemble des réseaux mis à jour et des cartes ou arbres de classification produits.

7 Bibliographie

- [1] EL HADDADI A., HATIM H., DOUSSET B., BERRADA I. & EL BAKKALI H. (2011). “*Reduce threats in competitive intelligence system: a generic information fusion access control model*”. International Journal of Database Management Systems, Academy & Industry Research Collaboration Center (AIRCC), Vol. 3 N. 1, p. 19-33, février 2011.
- [2] EL HADDADI A. (2011). “*Fouille multidimensionnelle sur les données textuelles visant à extraire les réseaux sociaux et sémantiques pour leur exploitation via la téléphonie mobile*”. Thèse de doctorat de l'université Paul Sabatier, Toulouse III, 09 décembre 2011.
- [3] FREEMAN, L.C. (1979). “*Centrality in social networks: Conceptual clarification*”. Social Networks, 6:223–258.
- [4] GAY B. (2011). “*Universal dynamics on complex networks, really? A comparison of two real-world networks that cross structural paths but ever so differently*”. Social Network Mining, Analysis and Research Trends: Techniques and Applications”, Eds, I-Hsien Ting, Tzung-Pei Hong and Leon S.L. Wang, National University of Kaohsiung, Taiwan, In: IGI Global, NY, USA. ISBN 978-1-61350-513-7 (hardcover) -- ISBN 978-1-61350-514-4 (ebook) -- ISBN 978-1-61350-515-1 (print & perpetual access).
- [5] GAY B. & LOUBIER E. (2009) “*Dynamics and evolution patterns of business networks*”. International conference on Advances in Social Network Analysis and Mining (ASONAM), Athens, Greece.
- [6] GAY B. & LOUBIER E. (2012) “*Worldwide Dynamic Evaluation of Complex Business Networks*”. Dynamic Analysis for Social Network. Editor: Carlos Andre Pinheiro, iConcept Press. ISBN 978-14610987-3-7.
- [7] HATIM H., EL HADDADI A., EL BAKKALI H. DOUSSET B. & BERRADA I. (2010). “*Approche générique de contrôle d'accès aux données et aux traitements dans une plateforme d'intelligence économique*”. VSST, Toulouse - France, 25 au 29 Octobre 2010
- [8] SALLES M. (2003). “*Modélisation des situations de décision dans une méthode d'ingénierie du besoin en I.E*”. Conférence IERA, Intelligence Economique : Recherches et Applications, Nancy, France, 2003.
- [9] SALLES M. (2005). “*De l'analyse du besoin des PME en IE à l'Intelligence Territoriale*”. Colloque Européen d'Intelligence Economique, Poitiers Futuroscope, ESCEM Poitiers, p. 414-427, Poitiers, France, 2005.

- [10] SOSSON D., VASSARD M. & DOUSSET B. (2001). "*Portail pour la navigation en ligne dans les analyses stratégiques, scientifiques et technologiques*". Veille stratégique, scientifique et technologique : VSST'2001, Barcelone - Espagne, 15/10/2001-19/10/2001, Vol. 1, p. 347-358, 2001.